The “Vertical distribution of ice nucleating particles over the boreal forest of Hyytiälä, Finland” by Brasseur et al. is an interesting study investigating how INPs vary with height in a region that could release a lot of INPs and is understudied. Aircraft measurements of INPs are particularly useful, because the majority of studies are ground-based and thus cannot usually provide measurements at the heights where clouds can form. The authors present both INP and aerosol data from several flights, which they use to normalize the spectra for insight into vertical efficiency. I especially thought the case study section was strong and gave convincing evidence for the possibility of periodic transport from the surface to the free troposphere but that most of the time it is different. It would be most interesting to investigate further (in the future) the potential mechanisms for the transport.

Despite the positives of this work, there are several things I believe can be improved upon, both in the figures and text. The main things I note are being clearer about the blank corrections, as they matter significantly for the free tropospheric samples; being extremely careful about not overstating claims: I don’t think some of the conclusions are supported by the data presented (as worded). This happens throughout the manuscript. I also thought the number of studies for comparing to previous work could be improved. In general, I think the manuscript reads too long, and could be tightened to improve its readability. I fully intend and hope my comments are helpful for the coauthors and offer a useful perspective. I would recommend it for publication, but only after these comments are given consideration and the claims in the text are represented better.

Response: We thank the referee for their feedback on the manuscript, which has greatly improved due to the referee comments and suggestions. Our responses are given below (in blue font) in the context of individual comments.

Specific comments:

Introduction: Should add Levin et al. (2019) as they made vertical measurements of INPs from the surface to free troposphere over California, for comparison. They found increased concentration with height.

Levin, E. J. T., DeMott, P. J., Suski, K. J., Boose, Y., Hill, T. C. J., McCluskey, C. S., Schill, G. P., Rocci, K., AlMashat, H., Kristensen, L. J., Cornwell, G. C., Prather, K. A., Tomlinson, J. M., Mei, F., Hubbe, J., Pekour, M. S., Sullivan, R. J., Leung, L. R., and Kreidenweis, S. M.: Characteristics of ice nucleating particles in and around California winter storms, J. Geophys. Res.-Atmos., 124, 11530–11551, https://doi.org/10.1029/2019JD030831, 2019.

Response: Thank you for pointing out missing literature. We propose adding the following text line 74:

“Similarly, Levin et al. (2019) observed an increase in INP concentration and in the fraction of total aerosol particles capable of ice nucleation from the surface up to approximately 7 km above sea level in wintertime in California, and suggested that pollution aerosols near the surface were poor sources of INPs.”

Lines 103-112: Much of this should go in the methods and is distracting from the main message of the introduction.

Response: We agree with the referee that some of the details might be distracting from the main message of the introduction. However, we also believe that the summary of the previous findings from the HyICE-2018 campaign and their link to this manuscript belongs more in the introduction than in the methods. Therefore, we suggest rephrasing the paragraph lines 103-112 to:

"In this study, we present filter-based measurements of INPs conducted at ground level and aloft in the boundary layer and free troposphere (up to an altitude of 3.5 km) in and above a Finnish boreal forest during spring 2018. The measurements were organized in the framework of a larger ice nucleation measurement campaign, called HyICE-2018, which took place at the Station for Measuring Ecosystem–Atmosphere Relations (SMEAR II; Hari and Kulmala, 2005) in Hyytiälä, Finland and presented in details in Brasseur et al. (2022). Results from the HyICE-2018 are also available from Paramonov et al. (2020), who presented ground-based INP concentrations measured with the Portable Ice Nucleation Chamber (PINC) during the first part of the campaign. In addition, the study from Schneider et al. (2021) extended their measurements for more than one year after the HyICE-2018 campaign and focused on immersion freezing INPs measured with the Ice Nucleation Spectrometer of the Karlsruhe Institute of Technology (INSEKT). They showed that the surface INP concentrations have a clear seasonal cycle that appears linked to the abundance of boreal biogenic aerosol. Building from these previously published results, the objective of this study is to describe the vertical variability in INP concentrations from ground level to the free troposphere above the Finnish boreal forest environment. To do so, we use instrumentation installed both onboard a measurement aircraft and at the SMEAR II measurement site, which allows for comparison between INP measurements and simultaneous measurements of many particle and meteorological variables."

Figure 1a: Map is hard to read, but I like the colored flight track. Can you improve the resolution (and maybe spatial extent)?

Response: Based on the Referees' comments, Fig. 1a was updated to a 3D plot to highlight the flight tracks more clearly. A map depicting the measurement site location in northern Europe is added to the Appendix. The updated figures are attached at the end of this document.

Line 160: A place with the basics of filter collection is needed: how many of each locational type were collected?

Response: We propose adding the following information line 199:

"In total, from the 19 measurement flights, 18, 16, and 16 filters were collected in the boundary layer, in the free troposphere, and at ground level, respectively."

Line 162: Please specify the length of soaking in 10% H_2O_2 and number of water rinses

Response: We added the missing information line 162:

“Before sampling, the filters were pre-cleaned by soaking them with 10 % H_2O_2 for 10 minutes. Afterwards, they were rinsed three times with deionized water that was passed through a 0.1 µm Whatman syringe filter.”

Line 193/Figure A1: Were the blank samples corrected in INPs/mL suspension space? I don’t think that is the unit you want to correct in if your ground samples were resuspended in 8 mL and boundary layer/free troposphere samples were resuspended in 5 mL, as that number is a function of resuspension volume. I think it is necessary to correct in total INPs/filter (multiply by resuspension volume), unless some blanks were resuspended in 8 mL to correct for the ground samples, and other blanks were resuspended in 5 mL. However, if everything was resuspended in 5 mL for your data, it would be ok and please make that clearer.

It would be helpful to indicate on Figure A1 which samples you ignored from being within your threshold, and you could also indicate in the text the percentage of filters you were able to keep. In the text, I would suggest providing more detail about the background corrections, and not only refer to previous literature (even though I know it is a related study). The reason is that often free troposphere/boundary layer airborne filters as you know are very close to the limit of detection, and so blanks can play a very big role in the answers and thus the conclusions that are drawn. Additional questions that I have, for example, are did you average the blanks and create a regression to subtract? For the samples that you adjusted because they were within a factor of two of the blanks, were you able to keep some points or did you remove the entire spectra? It would be great to indicate the points on a plot (Figure A1 for example) that were measured on the INSEKT but did not pass your criteria. I think it would be worthwhile to spend time to make this clearer. I think the criteria is acceptable, but at this point it is not repeatable.

Response: We thank the referee for their useful comment on the background correction. Both the samples and the handling blanks from the aircraft were suspended in 5 ml, thus the correction was done in the unit of INPs per ml. We modified the information line 189 to improve clarity:

“For this reason, and to enhance the INP content in the sample solution, the volume of filtered nanopure water was reduced from 8 to 5 ml for the samples collected onboard the aircraft.”

And lines 193-197:

“The INP concentrations reported here have been corrected for the background freezing levels of filtered nanopure water. The INP concentrations extracted from the aircraft samples were further corrected for the INP concentration derived from handling blank filters, which were collected onboard the aircraft without ambient air flowing through the membranes. Then, as the INP concentrations measured from the aircraft filters were rather low and close to the background signal derived from the handling blank filters (Fig. A2a), only the INP concentrations that were at least twice as high as the average background INP concentrations were considered significant and used in this study. More information concerning the handling blank correction can be found in Appendix A1. The INP concentrations extracted from the ground-level samples were well above the INP concentration derived from ground-level handling blank filters (Fig. A2b) and were therefore not corrected further.”

We also add a short section in the Appendix to describe the procedure followed for the background correction:

"The INP concentrations extracted from the aircraft samples were corrected for the INP concentration derived from handling blank filters collected onboard the aircraft. To do so, the INP temperature spectra obtained from the handling blanks were fitted exponentially and averaged to produce a single background curve used for background subtraction (black line in Fig. A2a). Only the INP concentrations that were at least twice as high as the background curve were considered significant and used in this study. This means that, for a given sample, only the data points meeting this criterion were used, while the data points not meeting the threshold were removed from the analysis (Fig. A2a)."

Finally, we updated Fig. A2a to include the single background and the data points that were removed. We added Fig. A2b which shows the ground-level data compared to the ground-level handling blanks, and we moved the previous Fig. A2 to Fig. A2c to regroup all the subplots together. The updated figure A2 is attached at the end of this document.

Figure 2: This is a very nice figure and is helpful to a general audience.

Response: We thank the referee for their positive feedback.

Paragraph beginning with Line 302: The ice onset is not helpful here because as you state, the volumes are very different. I would remove it, or if you keep it in, also qualify the free troposphere portion at the end as the volumes were the shortest (I understood 7 LPM for 1 hour average).

Response: We would like to keep the mention of the ice onset and thus added more information concerning the free troposphere samples as suggested by the referee:

Line 318: "[…] They also have an ice onset temperature colder than any other measurements shown in this study (approximately 4.5°C, 7°C and 10°C colder than the ice onset temperatures of the boundary-layer, ground-level and 24-hour measurements from Schneider et al. (2021), respectively). As mentioned previously, this is likely due to shorter sampling times used for the free troposphere samples (≈ 60 minutes)."

Figure 4: I would suggest you have a criteria for plotting/not plotting the histograms (at the cold end) based on the number of observations (e.g. >50%). I understand and appreciate you including the number of observations, but as most apparent in the ground histograms, the colder observations will be biased low based upon the more concentrated samples needing more dilution.

Response: We would prefer showing the entire dataset and not removing any data. We agree with the referee that some of the colder observations are biased low due to the limited number of data points at these temperatures, which is why we included the number of observations in Fig. 4c. To avoid confusion related to this matter, we propose mentioning this in the main text, line 305:

"Note that representing the data in this way might introduce some bias when the number of observations used to calculate the boxplots is more limited, for example at colder temperatures where some of the ground-level INP concentrations appear to be decreasing with decreasing temperature. The number of observations for each sample type as a function of temperature is highlighted in Fig. 4c."

Lines 334-337: The fact that the activated fraction brings the free troposphere closer to the rest of the observations than INP concentrations alone, would there be an argument for that making them more efficient (relatively speaking: still "less efficient" overall) as now some of the histograms overlap (especially with the previous study) over the range with many observations? It is unexpected to me that they are closer, and is an important finding to highlight more, even if it is not within the main message of the paper.

Response: We agree with the referee that this is an important observation that should be highlighted. We propose rephrasing lines 334-337 to:

"Figure 4b shows that there is more overlap between the activated fraction from all sample types compared to the INP concentrations shown in Fig. 4a. The activated fraction from the ground-level and boundary-layer samples are within the same order of magnitude, while the activated fraction from the free troposphere samples is overall lower, with some overlap with the ground-level samples below -20 °C. This suggests that, even though particles sampled in the free troposphere are overall less efficient INPs, there are a few cases where the free tropospheric INPs are as efficient as those sampled at ground level. These specific cases are further discussed in section 3.7."

Line 376-379: How do you reconcile the similarity of the activated fractions of the free troposphere to the Schneider et al. study (especially below -20) with this statement? INP concentration speaking, I agree with your statement. Lines 382-383 are affected as well.

Response: The activated fraction from the free troposphere samples is brought closer to the rest of the observations due to the few cases where the INP concentration (and activated fraction) observed in the free troposphere was higher. These cases are discussed in section 3.7, where we show that the free troposphere can sometimes be influenced by the boundary layer (itself influenced by the surface). We propose rephrasing lines 382-383 to:

"On the other hand, the lower INP concentrations measured in the free troposphere are most likely due to the lower particle concentrations encountered there combined with the fact that the free tropospheric particles might be less efficient INPs, as suggested by the overall lower activated fraction (Fig. 4b). In addition, the differences observed in the size distribution pattern suggest that the aerosol populations present in the free troposphere are different than those encountered in the boundary layer and at ground level. It is likely that these particles, and thus the INPs, were transported from distant regions via long-range transport, as discussed in section 3.4. Nevertheless, as mentioned previously, the activated fraction from the free troposphere samples sometimes overlap with the rest of the observations, in particular with the ground-level measurements from Schneider et al. (2021) and at temperatures below -20 °C. This shows that there are some cases where the activated fraction of the free troposphere samples is higher and similar to those observed at ground level. Such observation raises the question of whether surface particles might influence the free troposphere locally in some way. This question is further investigated in section 3.7, where we focus on the flights with the highest INP concentrations observed in the free troposphere."

Paragraph starting with Line 417: This paragraph is wordy and doesn't really present anything new. A point of needing more samples would be sufficient in the conclusions. I would suggest to trim/remove this. I do like the HYSPLIT analysis in this general section.

Response: We agree with the referee and propose removing the paragraph and moving part of it to the conclusions.

Line 444: Define CFDC at first use

Response: We added the acronym definition at its first use line 444.

Lines 447-450: There is not a sufficient explanation on why the aircraft OPS data was not used for all aircraft samples, even if you are assuming the air is similar enough. I would at least include that representation in the supplemental information for transparency, maybe as an additional figure or column. Were the percentages any different?

Response: We propose modifying the text lines 447-450 to the following:

"To calculate the total number concentration of particles with diameters larger than 0.5 µm used in these two parameterizations, we used the SMEAR II APS data for the boundary-layer samples (Fig. 7c, and e). This choice was motivated by the similarities between the size distributions and particle concentrations measured at ground level and in the boundary layer, the fact that the boundary layer was well-mixed during the aircraft measurements, and to investigate if ground-level measurements can be used in parameterizations to predict INP concentrations observed aloft in the boundary layer. For comparison, the INP concentrations predicted by the DeMott et al. (2010) and the Tobo et al. (2013) parameterizations calculated using the aircraft OPS data is shown in Fig. A5. On the other hand, since the free troposphere was characterized by distinct size distributions and particle concentrations, we used the aircraft OPS data to calculate the total number concentration of particles with diameters larger than 0.5 µm in the free troposphere (Fig. 7d, f, and g). For both sample types, the particle concentration was averaged over the sampling time of each sample."

With the following figure in the Appendix:

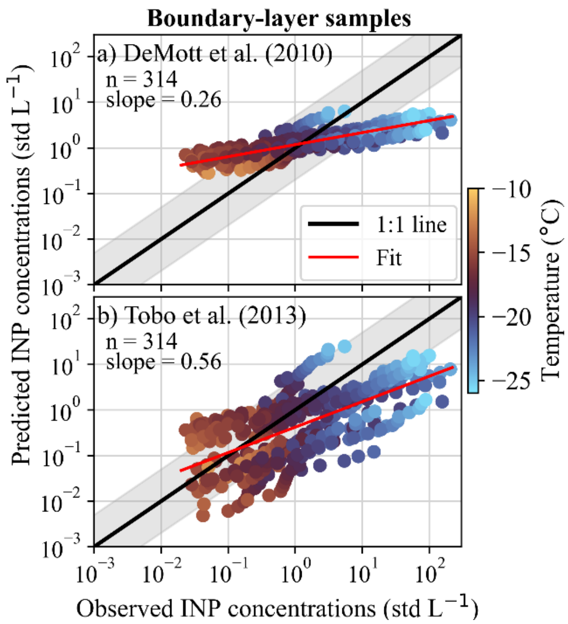


Figure A5. Comparison between the INP concentrations observed in the boundary layer and the INP concentrations predicted using the parameterizations from a) DeMott et al. (2010), and b) Tobo et al. (2013) using the aircraft OPS data. DeMott et al. (2010) reproduces 60 % and 32 % of the data points within a factor of 5 and 2, respectively. Tobo et al. (2013) reproduces 49 % and 17 % of the data points within a factor of 5 and 2, respectively. The black solid line represents the 1:1 line while the grey shaded area indicates a range of a factor of 5 from the 1:1 line. The red solid lines show a linear regression fit through the logarithmically transformed data points. The slope of the fit and the number of data points used is shown in each panel.

Finally, we propose mentioning the very similar percentages obtained line 475:

"Very similar results are obtained when calculating the DeMott et al. (2010) and Tobo et al. (2013) parameterizations using the aircraft OPS data instead of the SMEAR II APS data (Fig. A5), where DeMott et al. (2010) reproduces 60 % and 32 % of the data points within a factor of 5 and 2, respectively, while Tobo et al. (2013) reproduces 49 % and 17 % of the data points within a factor of 5 and 2, respectively. This highlights that both ground-level measurements and aircraft measurements from the boundary layer produce similar parameterization results, which suggests that ground-level measurements are sufficient for predicting INP concentrations aloft in the boundary layer."

Line 476: I think successfully is too strong of a word here. I agree that it performs the best out of all tested parameterizations (which is a nice finding of your study), but you need to take care to qualify and not overstate your conclusion. It's possible that another parameterization out there may fit your data better. The fit line shows the limitations.

Response: We propose removing the word "successfully" and rephrasing to:

"[…] among the parameterizations tested here, the Schneider et al. (2021) parameterization performs best at predicting the concentration of INPs in the boundary layer above a Finnish boreal forest environment."

Line 482: This is another sentence that needs qualifying. Yes, strictly speaking, based upon the factor of 5 and 2 percentage of points (which come with uncertainty as INP measurements have large error bars), the parameterization works better for the boundary layer. But visually, comparing Figure 7a and 7b, they look similar. Your statement "Thus the Schneider et al. (2021) parameterization can successfully represent the well-mixed boundary layer, but not the more remote free troposphere where INPs can be more scarce and originate from distant sources," is not convincing to me as the percentages and fit slopes are too similar to warrant a statement this strong. The word "successfully" appears again in the conclusions and abstract.

Response: We agree with the referee that the statement should be clarified and we propose rephrasing line 482 to:

"Thus, the Schneider et al. (2021) parameterization performs relatively better at representing the well-mixed boundary layer rather than the more remote free troposphere where INPs can be scarce and originate from distant sources."

In addition, we propose replacing "successfully predicted" line 24 of the abstract and line 655 in the conclusions by "best predicted".

Line 526: The onset difference here mostly looks related to limit of detection.

Response: We propose adding the following line 526:

"[…] although the ice onsets of our measurements are approximately 4 °C colder, likely due to differences in the limit of detection."

Figure 8: I think it would be helpful to indicate which studies are from what zone: ground, boundary layer, or free troposphere. This could either be accomplished in the legend with text or by grouping markers in the figure. It is confusing the way it is presented both in the figure and in the text right now. You could also trim the cold end to make the measurements easier to read, as few of yours go much colder than -25 °C. Adding an additional or two free troposphere study would also add value, as it seems that portion is lacking. This would help strengthen or potentially modify your statement in Line 541 saying the free tropospheric measurements fall within the same range as previous measurements. It would be insightful to compare how the free troposphere stacks against other free troposphere studies. Some that may be of use (both ground/airborne studies) would be Conen et al. (2022: https://acp.copernicus.org/articles/22/3433/2022/acp-22-3433-2022.html) Lacher et al. (2018: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018JD028338) Barry et al. (2021: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JD033752), in addition to the Levin et al. (2019) study I mentioned previously. You don't need to add all of these studies, I just think that it would bolster your figure/argument.

Response: We thank the referee for their useful suggestions which have greatly improved the figure. We modified Fig. 8 by grouping the studies by zone (boundary layer, free troposphere, and others) based on color in the figure, and we highlighted these groups in the legend. We also added the studies from Levin et al. (2019), Conen et al. (2022), and Barry et al. (2021) to provide more comparison to our samples:

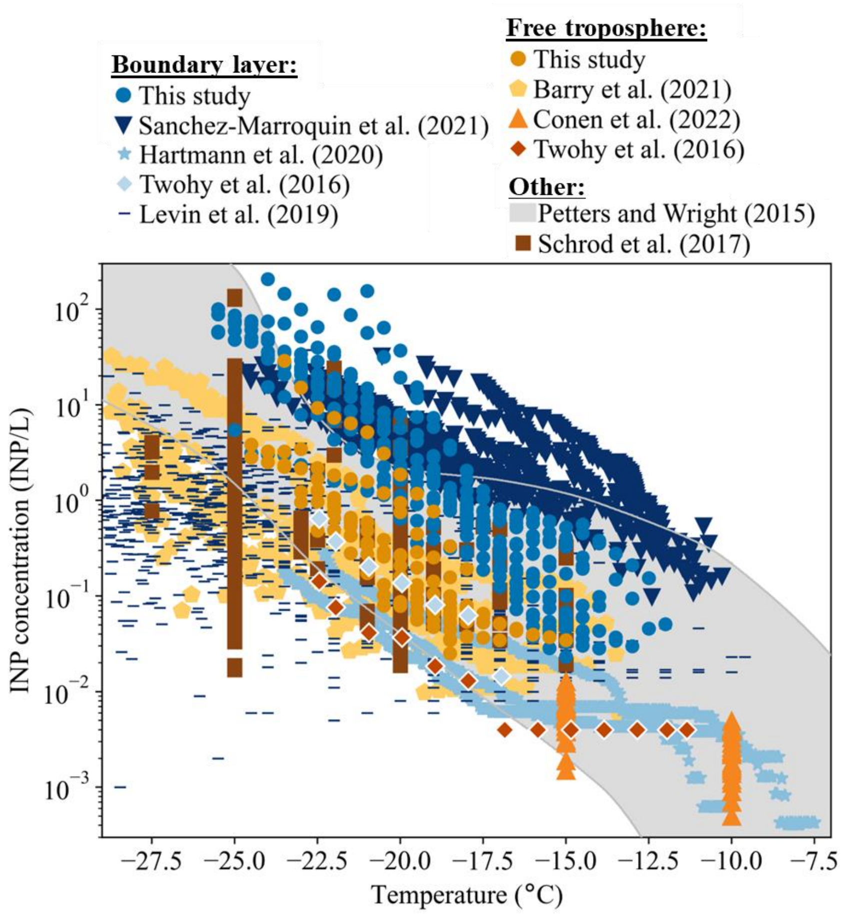


Figure 8: INP concentrations from the present study compared with literature data. The measurements from Sanchez-Marroquin et al. (2021) were conducted in the boundary layer in the South East of the British Isles, while Hartmann et al. (2020) conducted their measurements in the High Arctic boundary layer. For the Twohy et al. (2016) data collected in the western United States, the light blue diamonds represent a filter sampled within the boundary layer (at 1067 m a.g.l.), while the dark orange diamonds represent a filter sampled primarily in the free troposphere (between 897 and 3638 m a.g.l.). For the Levin et al. (2019) study conducted in California, the data represented here correspond to the measurements made in the boundary layer (below 2 km). The measurements from Barry et al. (2021) were conducted between 1300 and 5100 m above sea level in the western United States. The measurements from Conen et al. (2022) were conducted under free troposphere conditions at Jungfraujoch (3580 m above sea level) in the Swiss Alps. The grey band represents the data range given in Petters and Wright (2015) derived from precipitation samples collected around the world. The measurements from Schrod et al. (2017) were conducted between 500 and 2500 m a.g.l. (likely both in the boundary layer and in the free troposphere) over the Eastern Mediterranean.

We suggest rephrasing lines 524-546:

“Most of the data presented in this study fall within the mid-latitude data range given by Petters and Wright (2015; grey band in **Error! Reference source not found.**) derived from precipitation samples collected around the world, except for the highest INP concentrations measured between -18 and -24 °C in the boundary layer. Part of the data presented in this study also overlap with some of the INP concentrations reported by Schrod et al. (2017) who sampled Saharan dust plumes over the Eastern Mediterranean.

Concerning the boundary layer measurements, most of the INP concentrations presented in this study are higher than concentrations measured in the marine boundary layer in the Arctic during winter (Hartmann et al., 2020), in coastal California during wintertime (Levin et al., 2019) , and above a forested site in the western United States (Twohy et al., 2016). Compared to INP measurements conducted in the South East of the British Isles (Sanchez-Marroquin et al. 2021), the INP concentrations we observed in the boundary layer are about one order of magnitude lower for temperatures above approximately -18 °C, but are within the same order of magnitude for temperatures below approximately -18 °C.
The majority of the INP concentrations measured in the free troposphere in this study fall within the higher range of free tropospheric measurements from Barry et al. (2021) conducted during wildfire events in western US. The INP concentrations from Conen et al. (2022) observed under free troposphere conditions at Jungfraujoch in the Swiss Alps are relatively lower than the concentrations reported here at similar temperatures (≈ -15 °C). In addition, INP concentrations measured in the free troposphere over a forested site in the western United States (Twohy et al., 2016) are about one order of magnitude lower than the average INP concentrations observed in the free troposphere in the present study.
Thus, the INP concentrations measured in the boundary layer and in the free troposphere are mostly higher or within the same range as previous measurements from various regions. These observations illustrate that both the boundary layer and the free troposphere above the Finnish boreal forest are relatively rich in INPs with concentrations comparable to other environments."

Lines 557-559: Is it possible the water negatives were high in this particular sample, thus causing the gap between sample and negative to be too low, causing this junction? Sometimes you can just get unlucky and could also be physical if certain INPs are being inactivated in large volumes of dilution water. The shorter sampling time reason doesn't make a lot of sense, and you should be able to limit particle settling by resuspending the sample before dispensing. Ideally, I would suggest to rerun this sample because it is a part of your case study (maybe pick a lower level of dilution). If there isn't much sample left, you could just run the dilution. However, maybe there isn't any sample left in which case there's nothing you can do. In any case, I would remove the bit about shorter sampling time, because clearly you are still able to get detection to almost -15 °C.

Response: There is no remaining sample, and we are not able to re-analyze this sample. We propose rephrasing lines 557-559:

"Note that the discontinuity observed in the free troposphere sample from 17 May 2018 occurs at the dilution step. Possible explanations for this include the inactivation of INPs during dilution, the comparably low amount of sampled aerosol due to the shorter sampling time used for this filter (≈ 45 minutes), or insufficient redispersion of the suspension and the resulting inhomogeneity caused by particle settling (Harrison et al., 2018)."

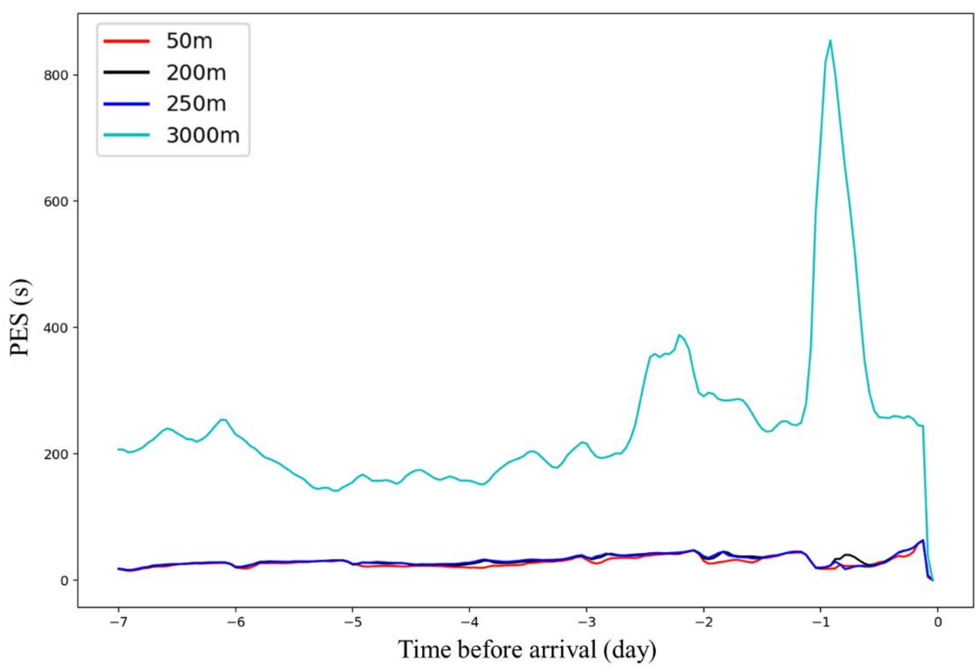
Section 3.7 General: Overall, I think the case study portion is well done and provides evidence at different angles. I really like the potential cloud processing depletion signal. The only caveat to this section I would mention is that just because the aerosol responds in a certain way doesn't mean the INP, a very small fraction of the total, will respond the same.

Response: We thank the referee for their positive feedback. We agree that we should mention the limitations of our analysis and we propose adding the following line 613:

"Furthermore, it is important to stress that, although the analysis of the particle number size distributions and concentrations gives valuable information on the vertical distribution and physical mixing state of the aerosol population, such information cannot necessarily be directly extended to the INPs, which represents a very small and highly variable fraction of the overall aerosol population (DeMott et al., 2010)."

Lines 628-629 and Figure 11c: I am not familiar with FLEXPART, but how certain can you be that there was little time below 200 m if the vertical resolution is 250 m? I say this because the lowest level seems to have an unexpected stripe with little variation. Could this be an artifact? It seems odd to me that there would be little surface influence the whole way. Again, I am no expert here, just pointing out an observation.

Response: The sentence line 629 should indeed be modified to 250 m a.g.l. instead of 200 m a.g.l. We updated the manuscript accordingly. The "stripe" visible at the lowest level is not an artifact but simply the result of the low PES values encountered in this layer, as highlighted in the figure below where the PES was simulated with a higher vertical resolution near the surface (for 50, 200 and 250 m) and compared with results obtained at 3000 m:



This figure clearly shows that the PES values at 250 m and below are substantially low compared to the much higher PES values at 3000 m a.g.l., as observed in Fig. 11c.

Line 630: The way it is written is confusing, I think it would be more correct to say that the air spent time in a particular layer.

Response: We propose rephrasing to:

“Results show that the air masses spent very little time below 250 m a.g.l., and therefore are less likely to have accumulated surface particles in transit. However, Fig. 11c shows that the airmasses did spend time in the boundary layer, even on the same day as the aircraft measurements.”

Conclusions: I think this section can be trimmed down and focused: it is lengthy and carries over some of the issues I noted in the main text.

Response: Based also on the comments from Referee 1, we recognize the need to recompose the conclusions. We propose a new text that is included in the response to Referee 1 and in the revised manuscript.

Lines 669-672: I don’t agree that you would need longer sampling times in order to do treatments on the suspensions: even 5 mL should give enough leftover volume to do one treatment and could be informative especially on some of your higher signal samples here. Definitely it would not be worthwhile for all of them, since some of them are near the negatives. It would be a good way to confirm the particles are similar with previous work. If you want to leave it for future work, that is fine, but I would suggest to remove/revise the explanation given.

Response: We agree with the referee that heat treatments would have been an important additional information in this study. We propose rephrasing the lines 669-672 to:

“Future measurements should include additional analysis of the chemical composition and heat sensitivity of the sampled INPs, in a similar manner to what Hartmann et al. (2020), Hill et al. (2016), and Sanchez-Marroquin et al. (2023) have done.”

Updated figures:

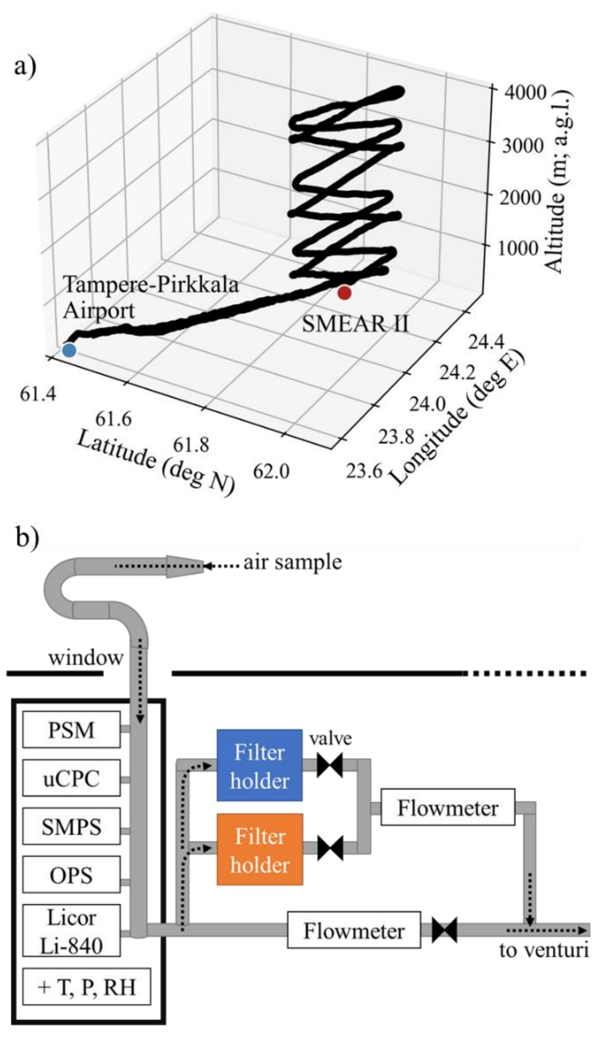


Figure 1: a) Example of a flight track from Tampere-Pirkkala airport to SMEAR II. The distance from the airport to the station is approximately 60 km. The location of SMEAR II with respect to Northern Europe is given in Fig. A1. b) Instrumental setup viewed from above inside the Cessna 172, described in detail in section 2.1.1.

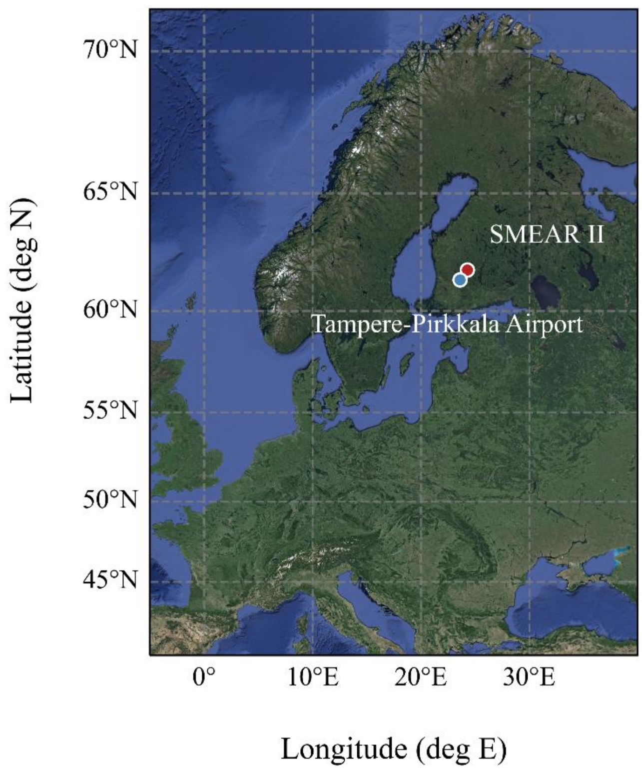


Figure A2: Location of SMEAR II and the Tampere-Pirkkala airport with respect to Northern Europe.

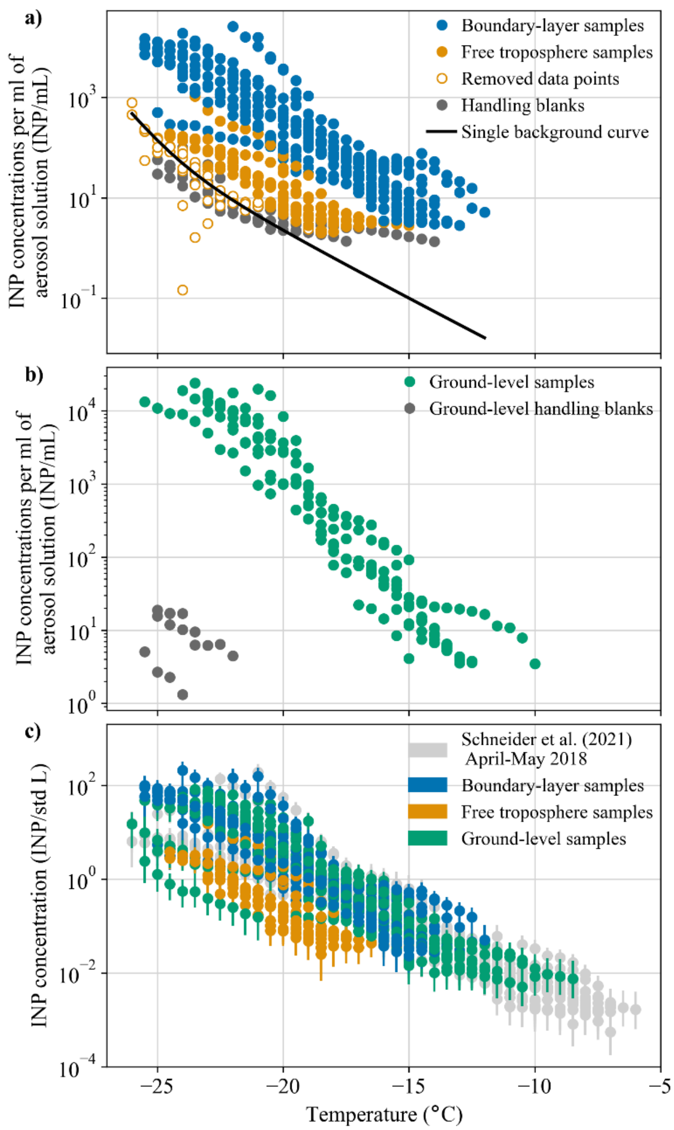


Figure A2: a) INP concentrations per ml of aerosol solution from the aircraft samples compared to the background signal derived from the handling blank filters collected onboard the aircraft. b) INP concentrations per ml of aerosol solution from the ground-level samples collected at SMEAR II at the same time as the aircraft samples compared to the ground-level handling blank filters. c) INP temperature spectra of all the samples collected during the aircraft measurement campaign together with the ground-level data from Schneider et al. (2021) collected in Hyytiälä during April and May 2018. The error bars represent the statistical as well as the systematic error of the INSEKT assay. More details related to the calculation of these error bars is given in Schneider et al. (2021).